Supporting Information

N-Acylethanolamine Acid Amidase (NAAA): Mechanism of Palmitoylethanolamide Hydrolysis Revealed by Mechanistic Simulations

Laura Scalvini,[1] Andrea Ghidini,[1] Alessio Lodola,[1] Donatella Callegari,[1] Silvia Rivara,[1] Daniele Piomelli,[2,3,4] and Marco Mor*[1]*

[1] Dipartimento di Scienze degli Alimenti e del Farmaco, Università degli Studi di Parma, Parco Area delle scienze 27/A, I- 43124 Parma, Italy

[2] Department of Anatomy and Neurobiology, University of California, Irvine, California 92697-4625

[3] Department of Pharmaceutical Sciences, University of California, Irvine, California 92697-4625

[4] Department of Biological Chemistry and Molecular Biology, University of California, Irvine, California 92697-4625

Corresponding Author

* Phone: +39 0521 905062. Fax: + 39 0521 905006. E-mail: alessio.lodola@unipr.it

* Phone: +39 0521 905059. Fax: + 39 0521 905006. E-mail: marco.mor@unipr.it

Table of Contents

1. NAAA-ARN19702 complexes: preparation and MD simulation with OPLS3e

Chains C and D taken from the X-ray structure of the human NAAA co-crystalized with the non-covalent inhibitor ARN19702 (PDB: 6DXX),[1] and corresponding to the α and β subunits, respectively, were used to model the enzyme-inhibitor complex. The structure was prepared using the protein preparation module of the Schrödinger 2018-2 suite.[2] Once hydrogen atoms were added, the orientation of hydroxyl groups and conformations of asparagine and glutamine side chains was adjusted in order to maximize the number of hydrogen bonds. The pH-activity profile of NAAA has a sharp positive peak within a pH interval ranging from 4 to 5.[3] While a first model was automatically generated by the protein preparation module of Maestro assuming an environmental pH of 4.5, the protonation state for the surroundings of the catalytic site was assigned on the basis of both the expected ionization state at acidic pH and on the influence of protonation on the dynamic stability of the crystallized structure.

The catalytic cysteine has two ionizable groups with different protonation and tautomeric states. The software PropKa,[4] based on empirical functions, estimates pKa values around 9.5 for both groups (Table S1). The thiol group was thus modelled in its neutral form, while for the amine group both states were

considered, given the proximity of the amine nitrogen to the cationic group of Arg300. Histidines and basic residues of NAAA were modelled in their protonated form. For aspartic and glutamic side chains within 10 Å of Cys126 (listed in Table S1), a combination of protonation states which stabilized the crystallized structure in MD simulations was searched.

Table S1. pKa values estimated with PropKa on 6DXX structure (Chain D)

Residue	**Estimated pKa value (PropKa*)**	**% Of protonated species at pH 4.5 (NH_3^+ / SH / COOH)**	**% Of unprotonated species at pH 4.5 (NH_2 / S^- / COO^-)**
C126 (amine)	9.55	99.999	0.001
C126 (thiol)	9.37	99.999	0.001
D145	2.82	2.047	97.953
D194	5.99	96.865	3.135
E195	7.76	99.945	0.055
D197	5.04	77.615	22.385
E297	4.14	30.387	69.613
D298	3.03	3.277	96.723
D299	2.52	1.036	98.964

**** Parse force field was used for calculations, setting environmental pH at 4.5.***

The side chain of Asp145 was modelled in its anionic form, which is consistent with the low pKa value estimated by PropKa, as it can accept a H-bond with the amine group of Cys126. The other acid residues lie on one of two flexible loops shaping the active site of NAAA, referred to as loop 1, carrying Glu297 and Asp299, and loop 2, carrying Asp194 and Asp197, respectively (Figure S1). For Glu195 PropKa estimates a high pKa value, which is consistent with the proximity to the carboxylate group of Asp197. For this reason, and because mutation of Glu195 to glutamine flattens the pH-activity peak, this residue was modelled in its neutral form. For the two aspartic acids, Asp298 and Asp299, PropKa estimates low pKa values, which suggests predominant anionic species. While Asp298 forms a H-bond with the guanidinium group of Arg, and for this reason was modelled as a carboxylate anion, Asp299 takes a H-bond with the alcoholic group of a threonine and we did not consider the low pKa value estimated by PropKa a sufficient reason to model it as a carboxylate. Thus, the protonation state of four acid residues

surrounding the catalytic state (Asp194, Asp197, Asp297 and Asp299) was considered uncertain, and its impact on the stability of the crystallized structure was assessed by MD simulations, building four systems (a-d in Table S2) where one of these carboxylic acids was deprotonated. For each proton configuration Cys126 was modelled in both its neutral (NH_2/SH) and positively charged form (NH_3^+/SH).

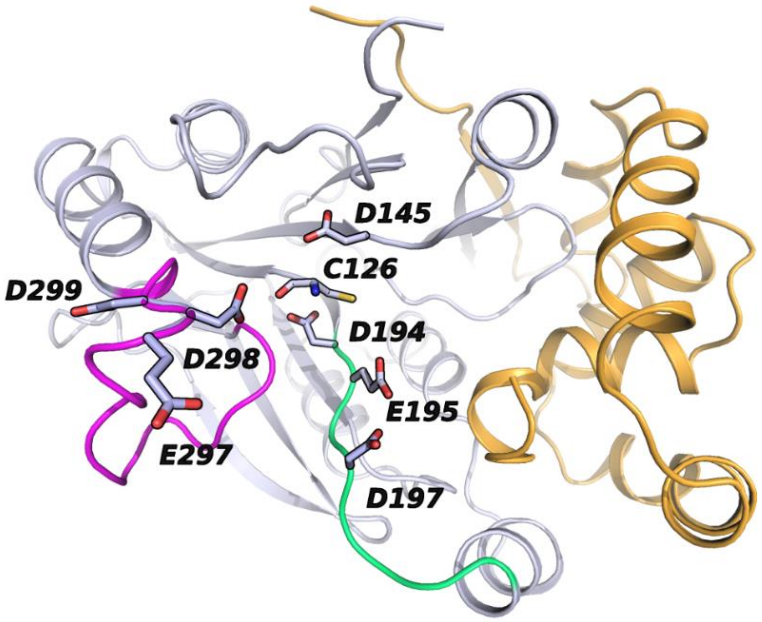


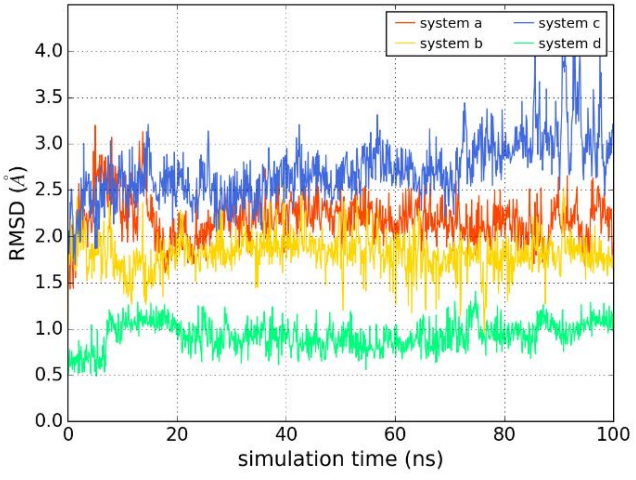
Figure S1. Representation of human NAAA structure (PDB structure 6DXX, in complex with ARN19702, not shown).[1] The active site is lined by two flexible loops (loop 1, residues 286-301, magenta cartoon; loop 2; residues 194-201, green cartoon) that carry Glu297 and Asp299, and Asp194 and Asp197, respectively.

Table S2. Protonic configurations of acid residues on the loops close to the active site.

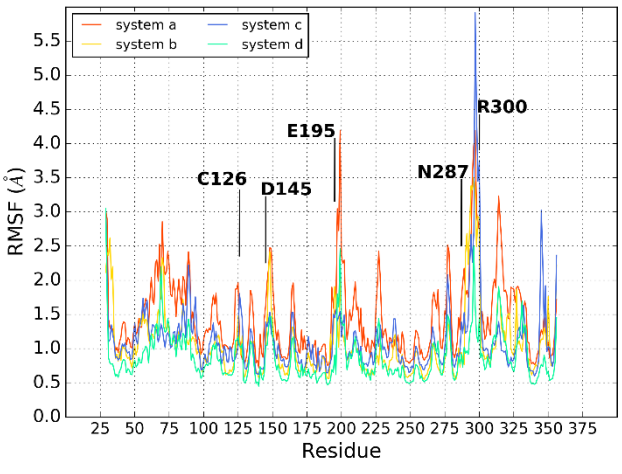
	System a	**System b**	**System c**	**System d**
D194	COO-	COOH	COOH	COOH
D197	COOH	COOH	COOH	COO-
E297	COOH	COOH	COO-	COOH
D299	COOH	COO-	COOH	COOH

The eight NAAA-ARN19702 complexes thus prepared were solvated by 14440 TIP3P water molecules and neutralized by adding an adequate number of Cl- ions (12 for the systems with neutral cysteine and 13 for the systems with cationic cysteine), setting the orthorhombic box dimensions to 10 Å. The molecule of Triton X-100 co-crystalized with NAAA was also included as part of the system. The resulting systems were then submitted to MD simulations with the OPLS3e force field.[5] These were carried out for 100 ns using Desmond 5.4[6] in NVT conditions at 298 K, after energy minimization of the system and equilibration by 15 ns in NVT conditions and 15 ns in NPT conditions. Bond lengths to hydrogen atoms were constrained by applying the M-SHAKE algorithm. Short-range electrostatic interactions were cut off at 9 Å, whereas long-range electrostatic interactions were treated using the smooth particle mesh Ewald method (PME). A RESPA integrator was used with a time step of 2 fs, while long-range electrostatic interactions were computed every 6 fs.

While the four systems with Cys126 modelled with a positive charge on the terminal amino group (NH_3^+/SH) were characterized by large fluctuations of the average root-mean squared deviation (RMSD) of the heavy atoms of residues lining the active site (data not shown), the simulations performed with Cys126 in its neutral form (NH_2/SH) were characterized by moderate RMSD (Figure S2A) and root-mean squared fluctuation (RMSF, Figure S2B), with system d showing the best profile.



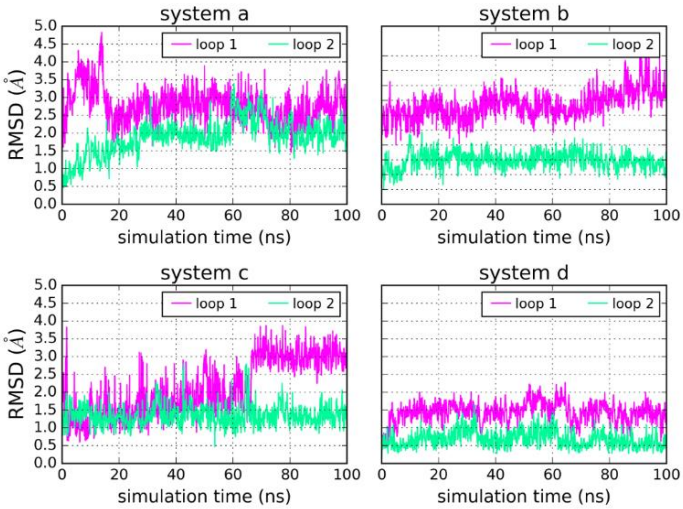
A



B

Figure S2. A) Analysis of the average RMSD of the heavy atoms of the residues shaping the active site of NAAA (Cys126, Asp145, Glu195, Asn287 and Arg300) during an MD simulation. The average RMSD was calculated for each of the 4 MD simulations performed with a different protonic configuration of the acid residues on loop 1 and 2 (systems a-d in Table S1). B) RMSF analysis for systems a-d taking the X-ray coordinates of the hNAAA-ARN19702 complex as reference structure.

These MD simulations proposed system d (characterized by Cys126, Asp194, Glu297 and Asp299 modelled in their neutral state and Asp197 modelled in its anionic form) as the most stable. Analysis of RMSD for the backbone heavy atoms of loop 1 (residues 286 to 301, magenta lines and cartoon in Figure S3) and loop 2 (residues 194 to 201, green lines and cartoon in Figure S3) showed that system d is characterised by higher stability for both loops, which maintain a stable conformation similar to that observed in the crystallographic structure. This result is only in partial agreement with the pKa values estimated by PropKa. While for Asp194 the neutral form is consistent with the estimated pKa value of 5.99, for Asp197 and Glu297 the anionic and neutral form, respectively assigned in system d, are estimated as 20-30% at pH 4.5, from the PropKa pKa values (Table S1). On the other hand, protonation of the acid group of Asp299 was essential for the stability of the catalytic site during MD simulation (system d *vs* a-c in Figure S2A), in spite of the low pKa value estimated by PropKa for this residue.



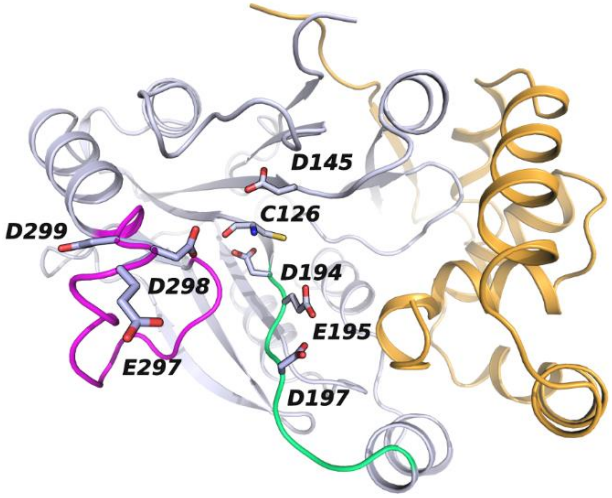


Figure S3. RMSD values for the backbone heavy atoms of loop 1 (residues 286-301, magenta cartoon) and loop 2 (residues 194-201, light green cartoon) during MD simulations on the different protonic configurations of residues Asp194, Asp197, Glu297, and Asp299.

Additional MD simulations were performed on the NAAA-ARN19702 complex with the AMBER force field to evaluate the dynamic stability of systems with alternative proton configurations for the residues directly involved in the catalytic process, i.e. Cys126, Asp145 and Glu195.

2. Definition of the QM region for NAAA-PEA complex

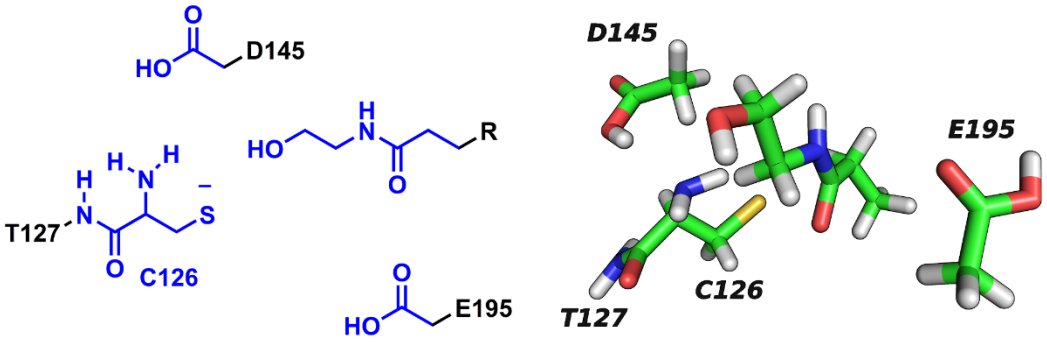
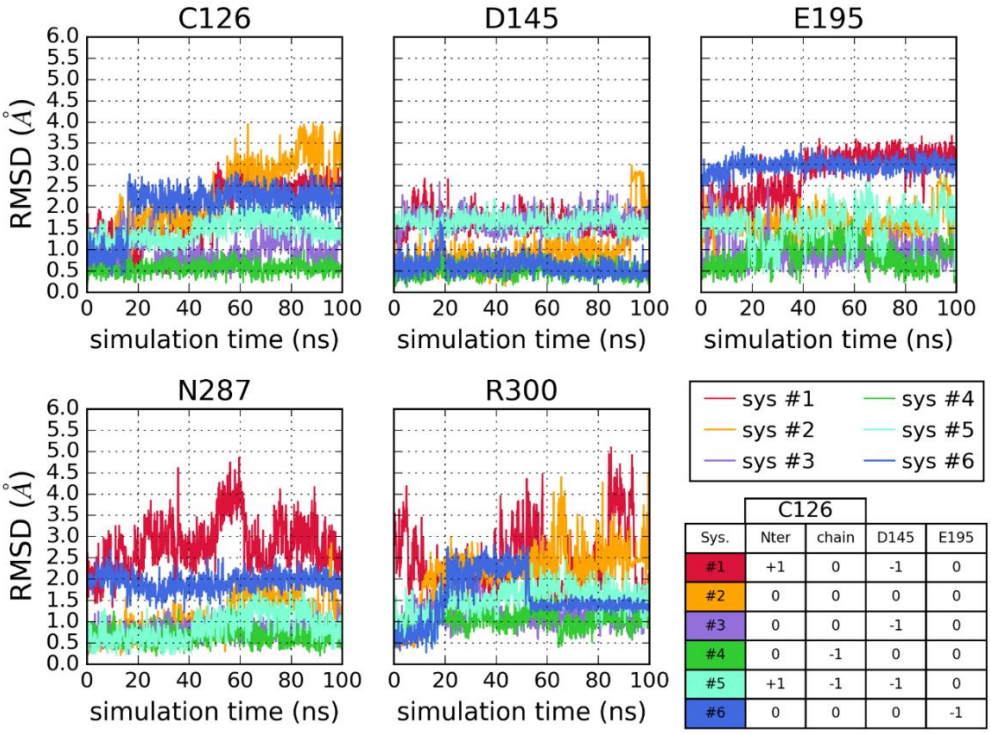


Figure S4. 2D and 3D representation of the atoms and of the link atoms included in the QM region

3. Stability of key residues of the active site of hNAAA in complex with ARN19702



	C126			
Sys.	Nter	chain	D145	E195
#1	+1	0	-1	0
#2	0	0	0	0
#3	0	0	-1	0
#4	0	-1	0	0
#5	+1	-1	-1	0
#6	0	0	0	-1

Figure S5. RMSD values for the heavy atoms of residues shaping the active site cleft of NAAA during an MD simulation. RMSD are calculated for 6 MD simulations performed starting each from a different protonic configuration of the active site (see also Figure 2).

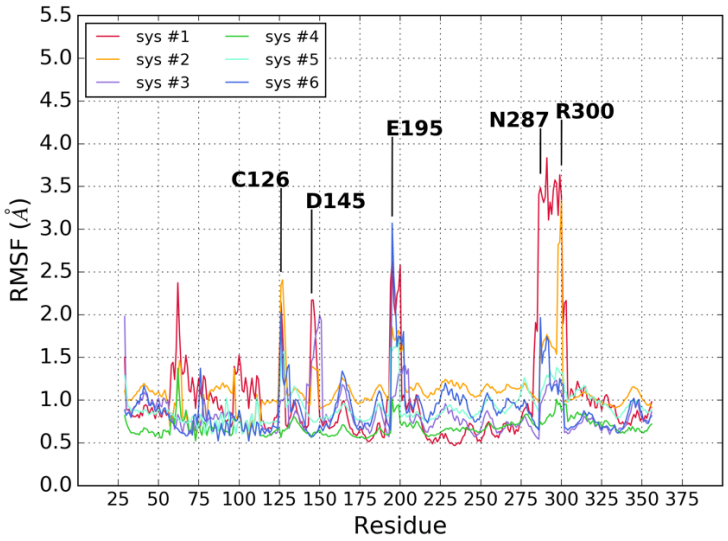


Figure S6. RMSF analysis for systems **#1**-**#6**. taking the X-ray coordinates of the hNAAA-ARN19702 complex as reference structure.

4. Superposition of PEA in its docked conformation on hNAAA on myristic acid in its crystalized conformation of ocNAAA.

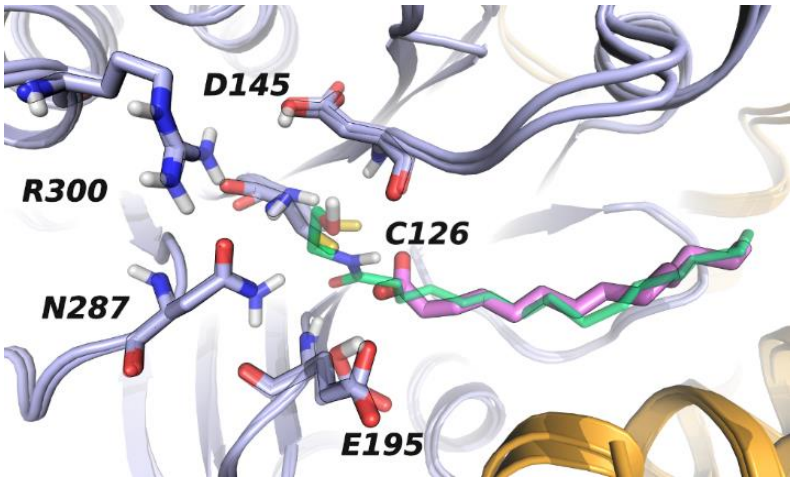


Figure S7. Superposition of the binding mode of PEA (green carbons) obtained by docking within hNAAA and the X-ray structure of ocNAAA in complex with myristic acid (pink carbons; PDB: 6DY1)

5. NAAA-PEA Michaelis complex: model building and equilibration

The NAAA-PEA Michaelis complex was built starting from a prepared model of the NAAA-ARN19702 X-ray structure (chains C and D), in which the proton configuration of NAAA corresponded to that of system **#4,** by docking PEA into the enzyme active site after inhibitor removal. This task was performed with Glide 7.9.[7] The docking grid was centered on the center of mass of residues Phe125, Cys126, Tyr146, Trp181, Glu195 and Asn287, setting the inner and outer cubic box dimensions to 12 Å and 20 Å, respectively. van der Waals radii of the protein and PEA atoms were reduced by a scaling factor of 0.8. The standard precision mode was used for pose-ranking. The best docking pose (G_{score}) was selected, and the resulting NAAA-PEA complex was imported in t-leap for parametrization with AMBER. The protein was modelled with the AMBERff15ipq force field,[8] while PEA and the Triton X-100 residue were parametrized using the GAFF.[9] The system was solvated by 16150 TIP3P water molecules and neutralized by adding 12 Cl^- ions. MD simulations were carried out for 100 ns in NVT conditions at 298 K, after the energy minimization of the system, and an equilibration of 15 ns in NVT and 15.5 ns in NPT

conditions, using the *pmemd* module of the AMBER16 software.[10] Likewise the MD simulations of the hNAAA-ARN19702 complex, electrostatic and van der Waals interactions were computed within a cutoff of 10 Å while long-range electrostatic interactions were treated using the PME approach. The covalent bonds involving hydrogen atoms were constrained with the SHAKE algorithm and a time-step of 2 fs was applied.

6. Interactions undertaken by the hydroxyl group of PEA during MD simulations

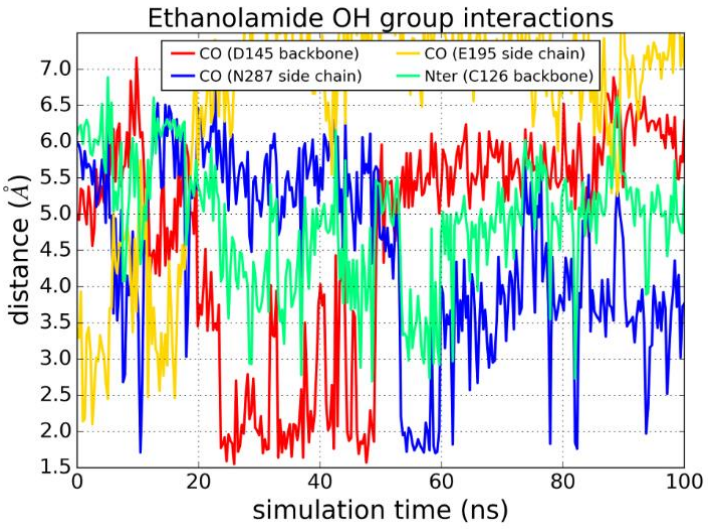


Figure S8. Analysis of the distances between the hydroxyl group of PEA and Asp145, Glu195, Asn287 and the N-terminal group of Cys126 throughout a 100 ns MD simulation of the Michaelis complex.

7. QM calculations on small models

A molecular model, consisting of Cys126 and of the acetic portion of Asp145, was used to compare the potential energy values obtained with DFTB3/3OB to those given by B3LYP-D3/6-31G*+ and MP2/6-31G*+. The optimized geometries taken from the proton transfer reaction were used to build three simplified models, which represented the protonation states corresponding to the configurations sampled in systems **#3**, **#4** and **#5**. Gas-phase energy calculations at B3LYP-D3/6-31G*+ and MP2/6-31G*+ levels were performed using Jaguar 10.0 software within the Schrodinger 2018-2 suite.[11] Default settings were applied.

8. Interatomic distances and Reaction coordinates for proton transfer PMF

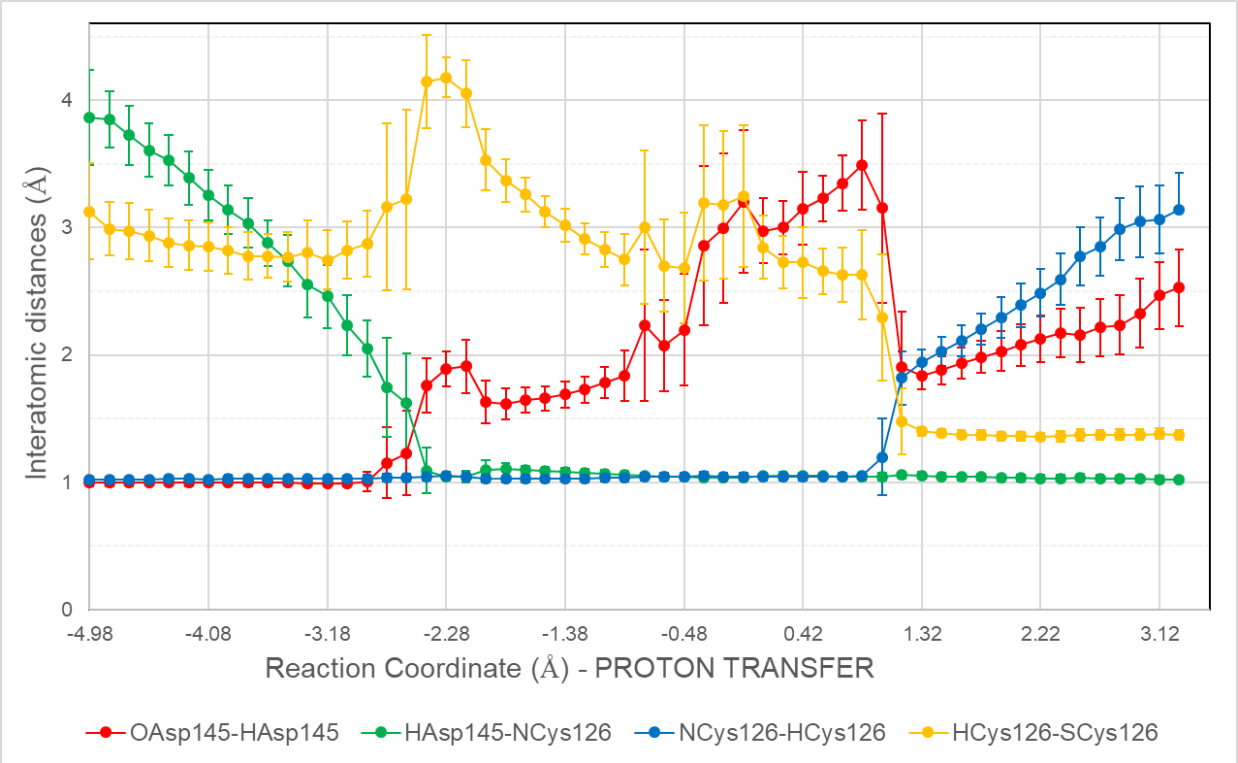


Figure S9. Average values of interatomic distances used for the definition of the reaction coordinate (RC) for the QM/MM US study of the transfer of a proton (H_{Asp145}) form the carboxylic oxygen of Asp145 (O_{Asp145}) to the amine nitrogen of Cys126 (N_{Cys126}), and of another proton (H_{Cys126}) from N_{Cys126} to the sulfur atom of Cys126 (S_{Cys126}). Starting from a dataset of 250 frames collected for each of the 56 US windows, points and error bars represent, the average values and standard deviations, respectively, of interatomic distances corresponding to intervals of RC values having width of 0.15 Å.

9. Potential Energy Surface (PES) for NAAA acylation (mechanism m1)

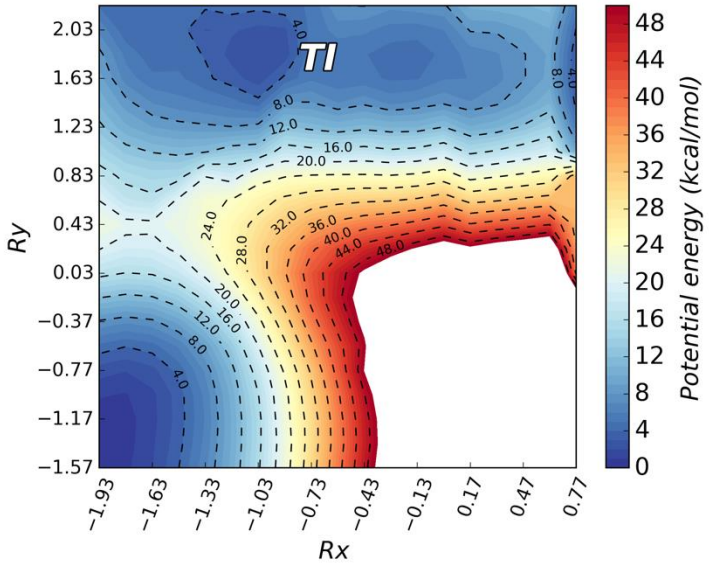


Figure S10. PES for NAAA acylation calculated at DFTB3/AMBER level. Potential energy values are given in kcal mol^{-1}, while reaction coordinates Rx and Ry are in Å.

10.Identification of the minimum free-energy path on computed FESs

The minimum free-energy paths on the free-energy surfaces reported for each calculation were determined using the minimum energy pathway analysis for energy landscapes (MEPSA) program.[12] Starting from a three-columns file reporting the free-energy values corresponding to each point of the map grid, the program samples the map and compares each point to the nearby points, checking if their value is equal, higher or lower. This procedure performed on the whole map allows the identification of the local minima (nodes). Following this step, the program finds the path connecting the origin point and the target point (corresponding to the reactants and the products, respectively) through an algorithm, named Global path generation method, that iteratively samples the map starting from the origin point and propagating toward the lowest energy points, until the target point is reached. The MEPSA records the iteration in which each point has been occupied, and once the target node is reached, a backward analysis from the target point to the origin is performed iteratively selecting the points with the lowest iteration

number. This algorithm, which is similar to Dijkstra's algorithm,[13] allows to identify the lowest energy path connecting nodes on the free-energy map.

11.Interatomic distance analysis during NAAA acylation

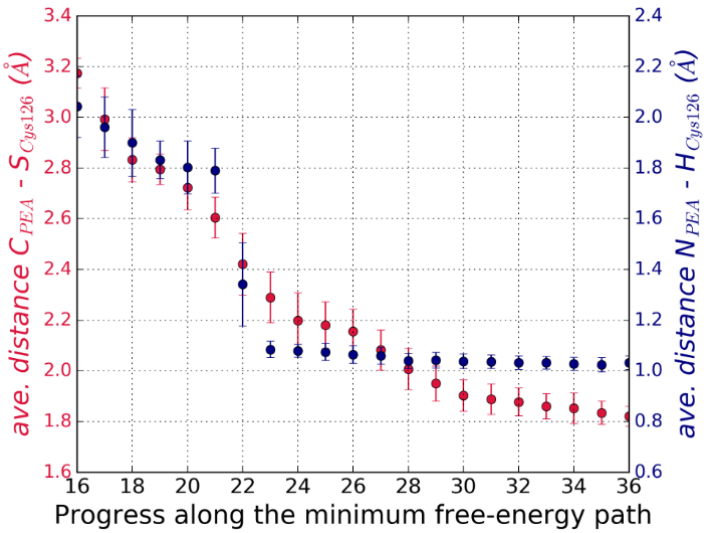


Figure S11. Analysis of the average distance between the S atom of Cys126 and the carbonyl carbon of PEA, and between the proton of the thiol group of Cys126 and the N atom of PEA throughout the umbrella simulations along a part of the 36 windows corresponding to the minimum free-energy path connecting the reactants to the products (36th window). While the proton transfer to the N atom of PEA is completed and irreversible after the 22nd window (corresponding to point *III* marked on the FES in Figure 7 in the main text) to the end of the acylation reaction, the decrease of the distance $S_{Cys126} - C_{PEA}$ shows a slower trend, suggesting that the acylation reaction requires the formation of an intermediate with the amide of PEA positively charged preceding the nucleophilic attack of Cys126 to the carbonyl carbon.

12.Indipendent replica for NAAA acylation with PEA (mechanism m1)

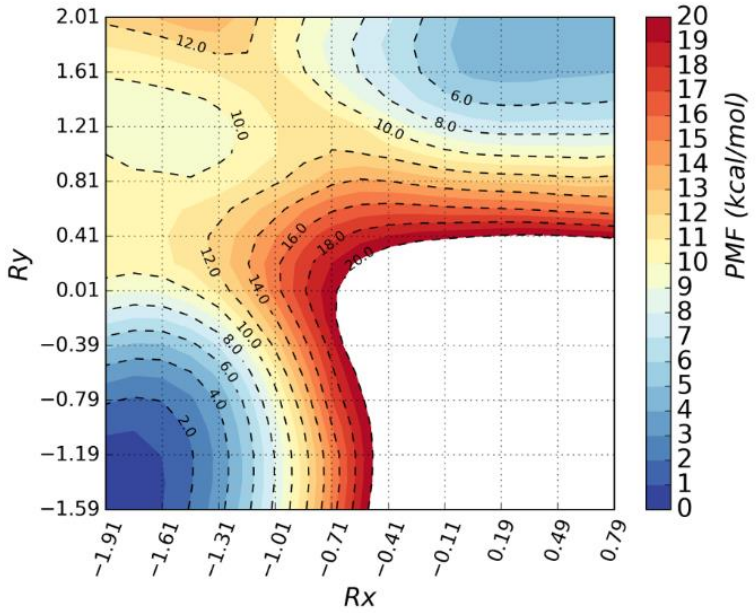


Figure S12. FES for NAAA acylation calculated at DFTB3/AMBER level. Free energy values are given in kcal mol^{-1}, while reaction coordinates Rx and Ry are in Å.

13.Role of Asp145 protonation in NAAA acylation with PEA

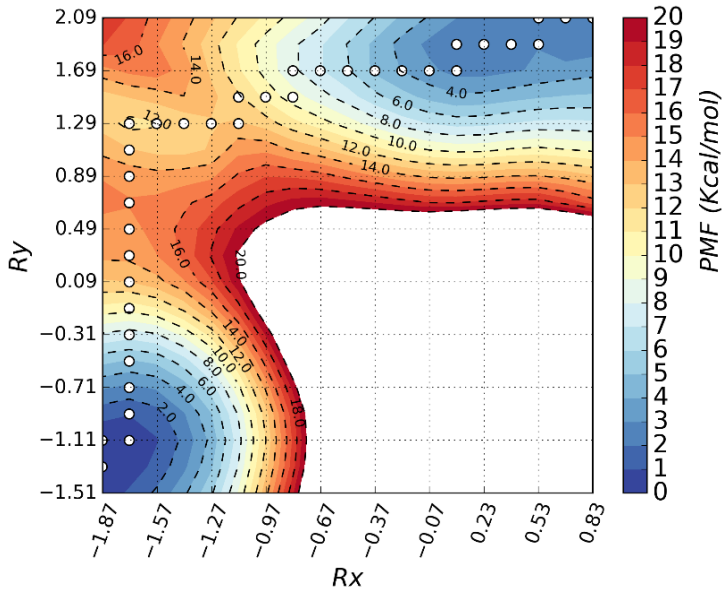


Figure S13. FES for NAAA acylation starting from configuration #**2**. Reaction coordinates Rx and Ry are expressed in Å. The process features a higher energy barrier compared to the one reported for

configuration #**3** (Figure 7 in the main text) for the two proton transfer events (from the Cys126 thiol group to the N terminal group, 15-16 kcal mol^{-1}, and from the N terminal group to the amide nitrogen atom of PEA, 12-13 kcal mol^{-1}).

14.Alternative mechanism: protonation of PEA amide by Cys126 (mechanism m2)

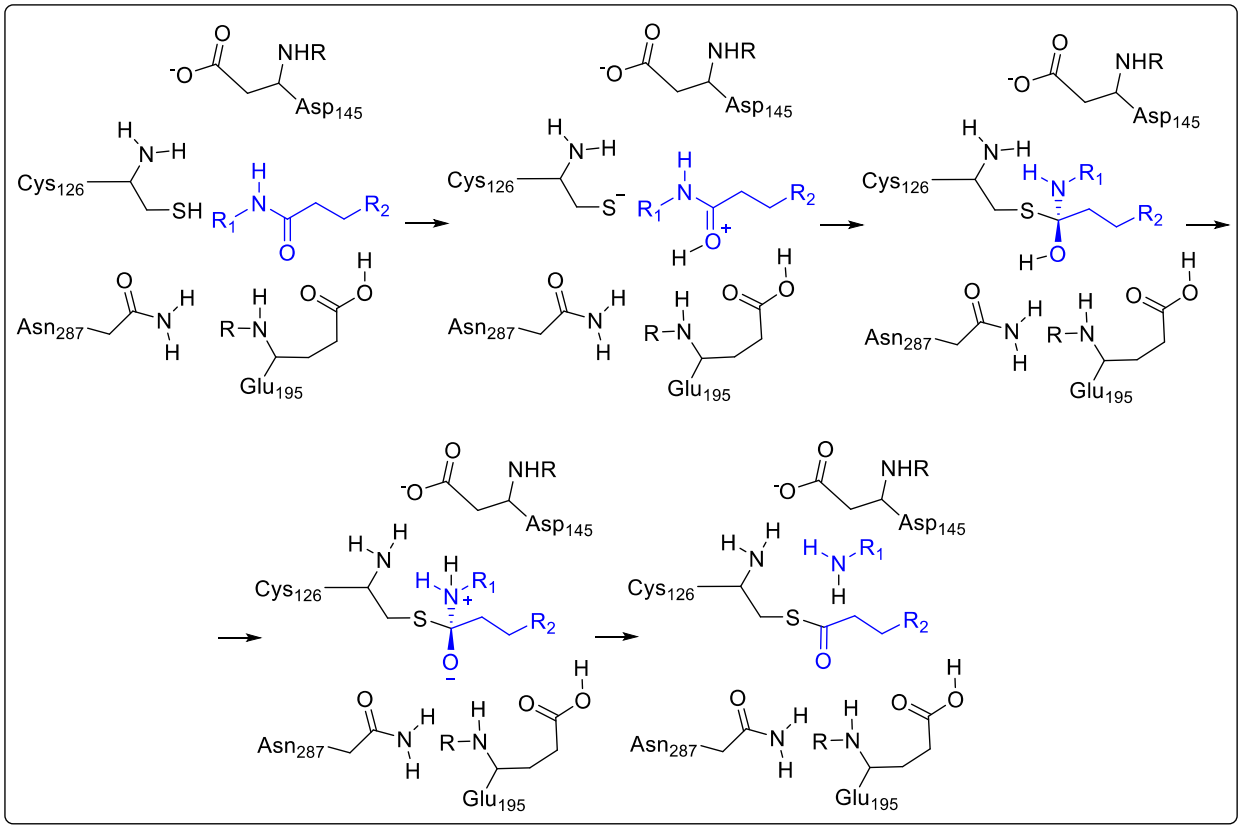


Figure S14. Alternative mechanism of NAAA acylation featuring the protonation of the PEA amide oxygen by the thiol group of Cys126. The ethanol chain at the nitrogen atom of PEA (R_1) and the complete chain of the palmitic acid moiety of PEA (R_2) are omitted for clarity.

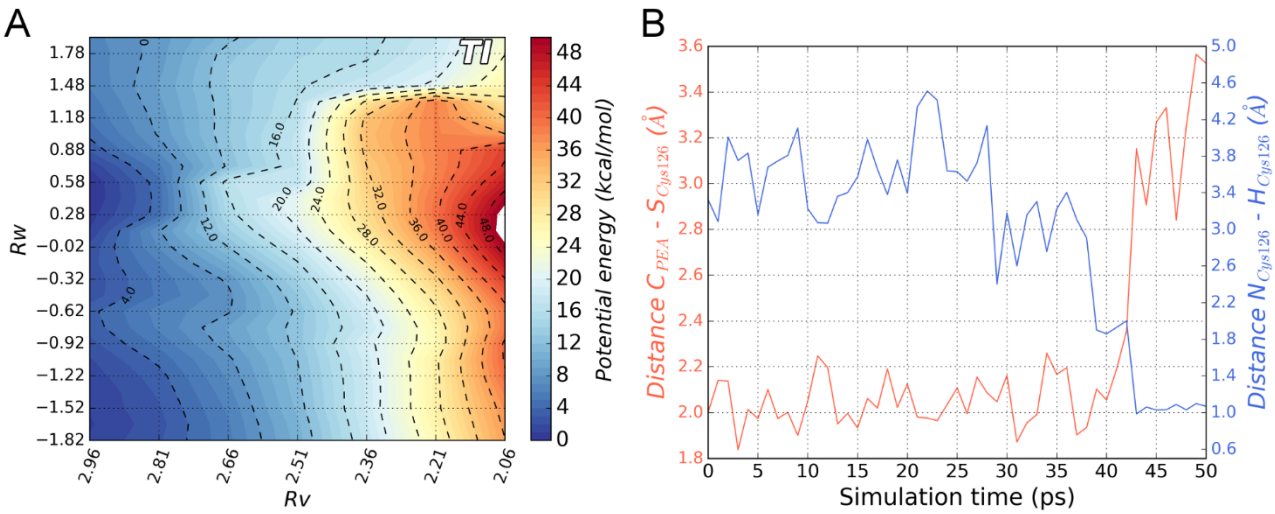


Figure S15. A) PES obtained along Rv (defined as the nucleophilic attack distance S_{Cys126}, C_{PEA}) and Rw (defined as the difference between the distance separating the nucleophilic sulfur from its acid proton − S_{Cys126}, H_{Cys126} − and the distance separating the same proton from PEA carbonyl oxygen − H_{Cys126}, O_{PEA}). The TI state is confined in the upper right corner of the PES and corresponds to a high energy region of approximatively 24 kcal mol^{-1}. B) Analysis of the distances between the sulfur of Cys126 and the carbonyl carbon of PEA (orange line) and between the N terminal group of Cys126 and the acid proton from the tetrahedral intermediate (blue line) throughout a 50 ps QM/MM MD simulation with restraints on the coordinate reaction values of the TI.

14.1. Technical details for mechanism m2

The same snapshot used to model the acylation reaction from configuration #**3** was selected as a starting point for the reaction leading to the formation of the tetrahedral intermediate. The structure was minimized, with SD for 500 steps and then with CG method, to an energy gradient of 0.005 kcal $mol^{-1}\cdot Å^{-1}$ at DFTB3/AMBER level, and then used as initial geometries for adiabatic mapping simulations at the same level of theory. The process was described using two reaction coordinates. The first one, Rv, described the nucleophilic attack and was defined as - [$C_{PEA}\cdots S_{Cys126}$]. The second coordinate, Rw, described the proton transfer from the nucleophile Cys126 to the carbonyl oxygen of PEA, and was

defined as the linear combination of the following distances: [$S_{Cys126}\cdots H_{Cys126}$] – [$H_{Cys126}\cdots O_{PEA}$]. The value of Rv ranged from 2.96 to 2.06 Å, with step size of 0.15 Å, while Rw ranged from -1.82 to 1.93 Å, with step size of 0.15 Å. At each point of the grid, a DFTB3/AMBER level minimization was performed to an energy gradient of 0.005 kcal $mol^{-1}\cdot Å^{-1}$ applying a harmonic restraint of 1000 kcal $mol^{-1}\cdot Å^{-2}$ on the reference value of the reaction coordinate. From this procedure, the potential energy surface reported in Figure S14A was obtained.

15.Alternative mechanism: protonation of PEA amide by Glu195 (mechanism m3)

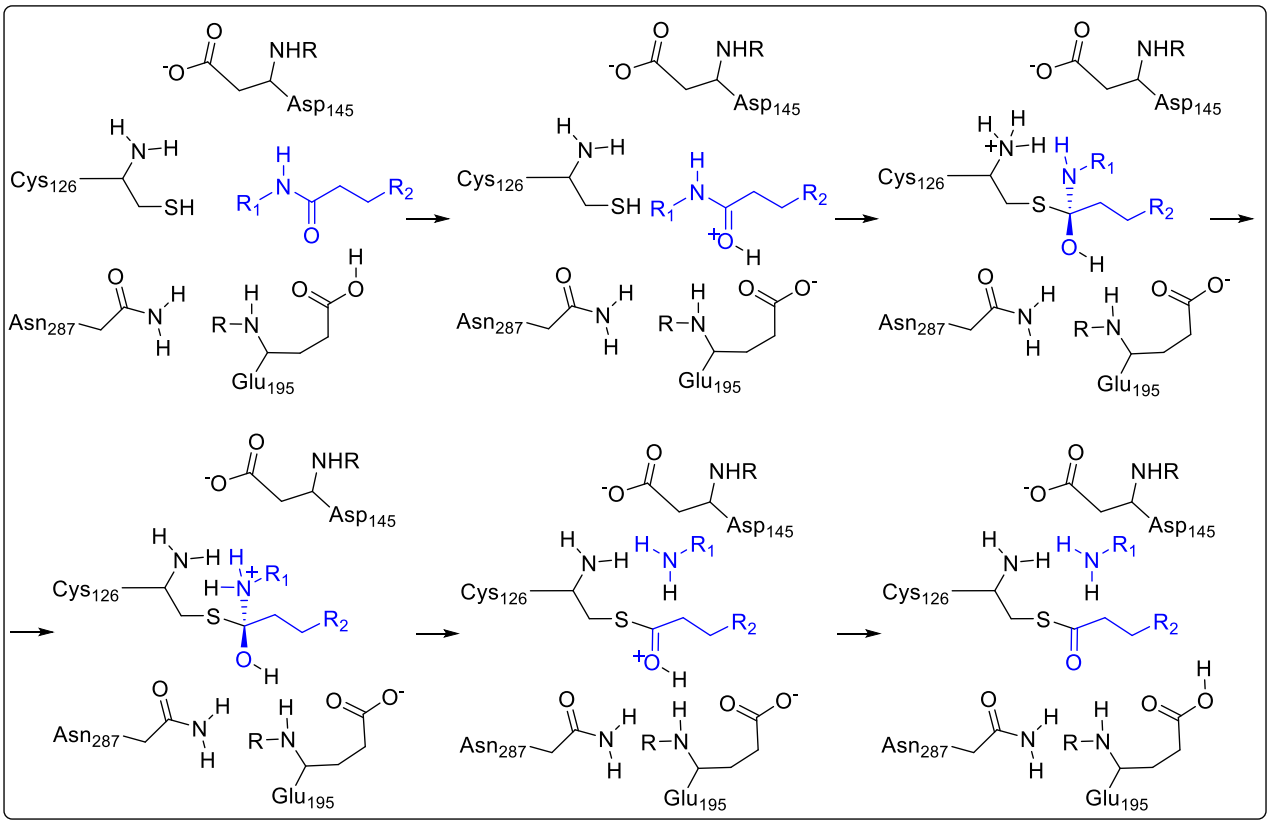


Figure S16. A possible mechanism of NAAA acylation occurring through protonation of the carbonyl oxygen of PEA by Glu195. The chain at the nitrogen atom of PEA (R_1) is omitted for clarity. For the same reason, the full chain of the palmitic acid moiety of PEA (R_2) is not displayed.

16. NAAA acylation (mechanism m1) by path-collective variable simulations

16.1. Technical details

A set of representative frames belonging to the minimum potential energy path for Cys126 acylation (taken from the PES of Figure S11) or from an arbitrary diagonal path featured by higher energy taken from the same PES of Figure S11, was used as the initial guess-path. In both cases, the selection was performed following the procedure described by Branduardi et al.[14] The C, O, and N atoms of the amide group in PEA and the N, Cα, Cβ, S and H atoms in Cys126 were used as reference atoms to describe the reaction path. Fifty frames, spaced by an average root-mean-squared deviation (RMSD) of 0.1 Å, represented the initial guess-path connecting the Michaelis complex to the acylenzyme. The collective variables *S* and *Z* were used to describe the position of the reference atoms (*R*) relative to the path. *S* and *Z* are defined as follows:

$$S(R) = \frac{\sum_{i=1}^{P} i e^{-\lambda(R-R(i))^2}}{\sum_{i=1}^{P} e^{-\lambda(R-R(i))^2}} \qquad \text{eq. S1}$$

$$Z(R) = -\frac{1}{\lambda} \ln \left(\sum_{i=1}^{P} e^{-\lambda(R-R(i))^2}\right) \qquad \text{eq. S2}$$

Where: i is an index representing the frame number within the path, ranging from 1 to P=50; $(R - R(i))^2$ is the squared deviation from the i-th frame of the path for the atoms of the reference set; λ is a smoothing parameter. The simulations were performed using the *sander* module of AMBER16 coupled with PLUMED 2.4.1.[15] The system was pulled along the collective variable S, starting from 1 (*i.e.* the Michaelis complex) and proceeding to 50 (*i.e.* the acylenzyme). A force constant of 300 kcal $mol^{-1} \cdot Å^{-2}$ was applied, and the overall simulation was performed at a velocity of 0.25 S steps/ps. To prevent the system to go exceedingly far from the initial path, Z was restrained to a value of 0.005 $Å^2$ by applying a force constant of 200 kcal $mol^{-1} \cdot Å^{-8}$. At the end of each steered-MD simulation, the frames collected were aligned to the first frame and interpolated using the Catmull-Rom splines method, in order to generate a larger set of structures from which to retrieve a smooth guess path.[14,16] An RMSD-based

distance matrix between each pair of frames was computed in order to extract a new set of 50 reference frames, equally separated from each other (0.1 Å), that was then employed as guess path for the following simulation run. This procedure was applied iteratively for 8 simulation runs, when the reaction path did not change anymore moving from a steered-MD simulation to the following one.

16.2. Feasibility study: comparison with US simulations for mechanism m1

To evaluate the feasibility of the PCV approach, we initially checked if the PCV method could give results comparable to those obtained with US simulations when modelling the reaction mechanism ***m1***. A set of configurations corresponding to the minimum potential energy path on the PES reported in Figure S10 was used as reference-reaction coordinate, or guess-path, and a series of steered-MD simulations was performed on the *S*/*Z* space. The iterative process that allows to update and optimize the guess-path was repeated until the path did not change passing from a simulation to a new one. Analysis of eight PCV simulations shows that the geometries explored along the path-collective variable *S* rapidly converge on geometries similar to those sampled by US simulations for acylation (Figure S17A). Accordingly, during the PCV simulations the values of the distances implemented in the definition of the reaction coordinates for US simulations fluctuate in the regions of Rx and Ry corresponding to the minimum-free energy valley of the US simulations (Figure S17A). Importantly, when we started PCV simulations from a guess-path built from geometries with higher free energy, such as those laying on the diagonal of the FES in Figure 7, the PCV approach quickly converged to a path similar to the minimum free-energy valley of US simulations (Figure S17B).

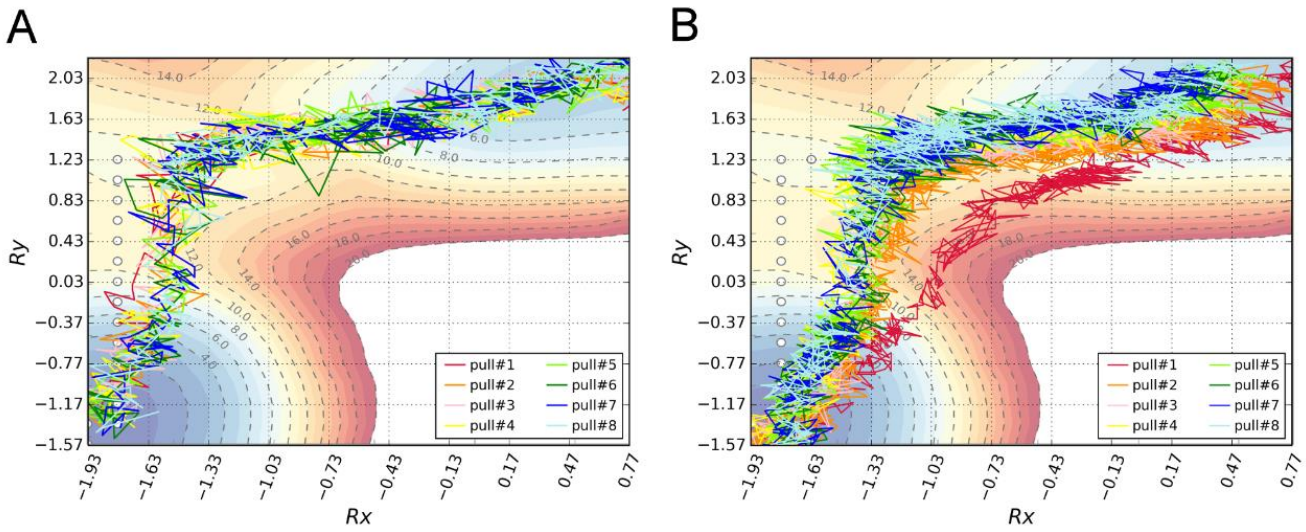


Figure S17. Exploration of the reaction paths for NAAA acylation by PCV approach. A) Geometries sampled by steered-MD runs projected on the Rx and Ry space. B) Geometries sampled by steered-MD runs projected on the Rx and Ry using a (high-energy) guess-path to connect reactants to products (see text). White dots represent the minimum free-energy path connecting reactants and products according to the QM/MM US simulations of Figure 7.

The shape of the work-profile calculated along *S* for the last run of steered-MD resembles to some extent the free-energy profile for NAAA acylation deduced from the FES reported in Figure 7 in the main text and projected on the S variable in Figure S17. The higher energy values of the work curves (compared to that obtained from US) depend on the non-equilibrium character of the chemical transformation performed by steered-MD.[17]

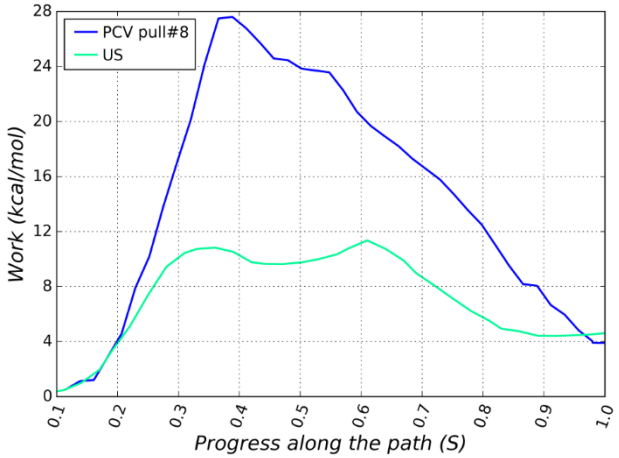


Figure S18. Free-energy profile (US sampling) and work curve (steered-MD) projected along *S*.

16.3. Computational cost for PCV and US simulations

The amount of resources spent for simulating NAAA acylation by PEA (mechanism m1) with US and PCV simulations has been evaluated in terms of core hours used, taking into account the:

- elapsed time (in hours) for each window (US) or run (PCV),
- the number of cores used for each window (US) or run (PCV),
- and the total number of windows (US) or runs (PCV) performed

The total amount of core hours used for US simulations is 57000 h (37.5 hours x 4 cores x 380 windows), while the number of core hours used for PCV calculations is lower (15872: 62 hours x 32 cores x 8 runs).

17.NAAA acylation (mechanism m3) by path-collective variable simulations

17.1. Technical details

For NAAA acylation involving protonation of PEA by Glu195 a preliminary guess path was generated through a set of consecutive steered-MD simulations. In the first step, protonation of the PEA carbonyl by the side chain hydrogen of Glu195 was modelled using the difference of the distances [$O_{Glu195}\cdots H_{Glu195}$], and [$H_{Glu195}\cdots O_{PEA}$]. The second step simulated the proton transfer from the thiol to the amine group of Cys126 and the nucleophilic attack of the thiolate to the carbonyl carbon, using the

distance combination [$S_{Cys126}\cdots H_{Cys126}$] – [$H_{Cys126}\cdots N_{Cys126}$] – [$S_{Cys126}\cdots C_{PEA}$] as reaction coordinate. The third step involved protonation of the leaving group, using the difference between the distances [$N_{Cys126}\cdots H_{Cys126}$] and [$H_{Cys126}\cdots N_{PEA}$] as reaction coordinate, and the fourth simulated the breakage of the bond [$N_{PEA}\cdots C_{PEA}$]. A final step consisted in the proton transfer from the protonated acylenzyme to Glu195, modelled using the difference of distances [$O_{PEA}\cdots H_{Glu195}$] and [$H_{Glu195}\cdots O_{Glu195}$] as reaction coordinate. In each step, the target value of the reaction coordinate was reached applying a force constant of 300 kcal $mol^{-1}\cdot Å^{-2}$ and setting the velocity to 0.001 Å/ps. The trajectories obtained from these five steps of steered-MD were merged and a set of representative frames, describing the entire reaction process, was extracted applying the procedure described in reference by Branduardi et al.[14] The C, O, and N atoms of the amide group of PEA and N, Cα, Cβ, S, H of Cys126 were used as reference atoms for the frame selection process. This procedure led to the extraction of 40 frames, spaced by an average RMSD of 0.1 Å. The system was pulled along S, starting from 1 (*i.e.* the Michaelis complex) and proceeding to 40 (*i.e.* the acylenzyme). A force constant of 300 kcal $mol^{-1}\cdot Å^{-2}$ was applied, and the overall simulation was performed at a velocity of 0.25 S steps/ps. In order to prevent the system to get exceedingly far from the initial path, Z was restrained to a value of 0.005 $Å^2$ by applying a force constant of 200 kcal $mol^{-1}\cdot Å^{-8}$ At the end of each steered-MD simulation, a new set of frames was extracted and employed as guess path. This procedure was repeated (8 consecutive runs) until the reaction path did not change anymore moving from a steered-MD simulation to the following one.

17.2. Evolution of key distances during the optimization process

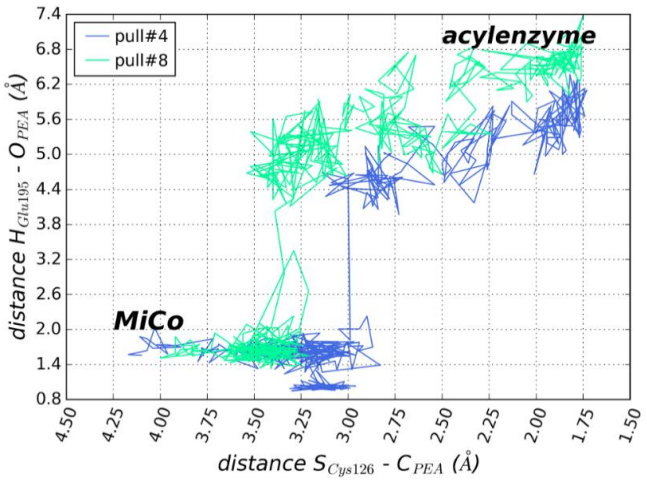


Figure S19. Analysis of key distances between the S atom of Cys126 and the carbonyl carbon of PEA, and between the acid hydrogen of Glu195 carboxylic group and the carbonyl O of PEA registered in two representative runs on *S* and *Z*. In pull #4 (blue line) the proton exchange from Glu195 to PEA still occurs (reduction of the $H_{Glu195} - O_{PEA}$ distance) as a reversible event, concluding before the formation of the $S_{Cys126} - C_{PEA}$ bond. In the eighth run (green line) Glu195 does not participate in any way to the catalysis, and the carboxylic acid of Glu195 changes conformation, pointing the acidic hydrogen far from the catalytic site.

18. Potential Energy Surface (PES) for NAAA deacylation

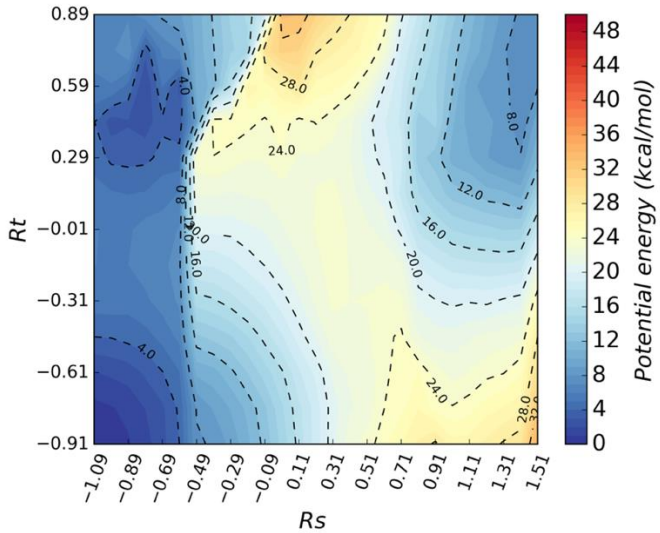


Figure S20. PES for NAAA acylation calculated at DFTB3/AMBER level. Potential energy values are given in kcal mol^{-1}, while reaction coordinates Rt and Rs are in Å.

19. Analysis of the interactions taken by the hydroxy group of PEA along the acylation reaction.

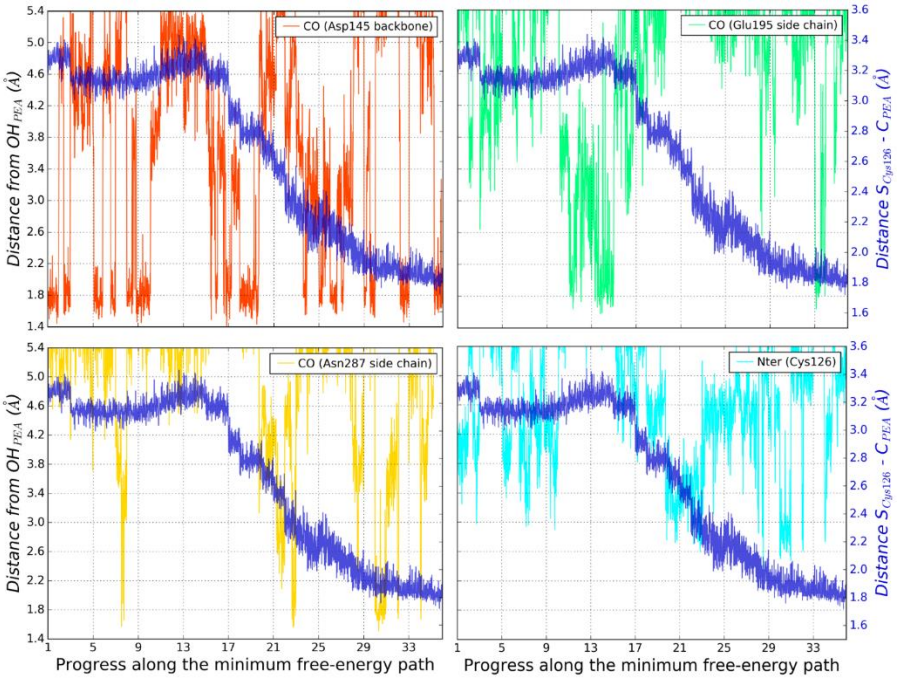


Figure S21. Analysis of distances between the H atom of the hydroxyl group of PEA and Asp145 (red line), Glu195 (green line), Asn287 (yellow line) and the N-terminal group of Cys126 (cyan) registered for the 36 US windows of the minimum free-energy path throughout the acylation process represented by the progressively lowering distance between the S atom of Cys126 and the carbonyl atom of PEA (blue line).

20. Preparation, MD simulation and QM/MM studies on the NAAA-PMA complex

20.1. Model Building: construction of NAAA-PMA Michaelis complex and MD simulation

The NAAA-palmitoyl methyl amide (PMA) complex was built starting from the chain C and D of the protein in complex with ARN19702 (PDB 6DXX),[1] modelled with the protonation state described in

system **#3**. The docking studies were performed using the Glide 7.9 software.18 The grid was centered on the center of mass of residues Phe125, Cys126, Tyr146, Trp181, Glu195 and Asn287, setting the inner and outer cubic box dimensions to 12 Å and 20 Å, respectively. The van der Waals radii of the protein and PMA atoms were reduced by a scaling factor of 0.8. The protein was modelled with the AMBERff15ipq force field,[8] while the PMA molecule was parametrized using the GAFF.[9] The molecule of Triton X-100 co-crystalized with NAAA was also included as part of the system and was parametrized using the general Amber force field. The system was solvated by approximately 16160 TIP3P water molecules and neutralized by adding 11 Cl- ions. A 100 ns MD simulation was carried out in NVT conditions at 298 K, after the energy minimization of the system, and an equilibration of 15 ns in NVT and 15.5 ns in NPT conditions, using the *pmemd* module of the AMBER16 software.[10] The electrostatic and van der Waals interactions were computed within a cutoff of 10 Å, while long-range electrostatic interactions were treated using the PME approach. The covalent bonds involving hydrogen atoms were constrained with the SHAKE algorithm and a time-step of 2 fs was applied. A representative frame at the end of the equilibration protocol was selected as starting point for a 400 ps QM/MM MD simulation. As with the NAAA-PEA complex, the whole Cys126, the backbone NH group of Thr127, the N-methylpropionamide fragment of PMA, and the acetic acid fragments of Asp145 and Glu195 were treated with the density functional tight-binding type 3 (DFTB3) approach. The resulting QM system was composed of 45 atoms including 4 link atoms.

3.2 Acylation (NAAA-PMA complex)

A snapshot taken from the 400 ps QM/MM MD simulation of the NAAA-PMA complex was selected as starting point for the acylation reaction. The structure was minimized, with steepest descent (SD) for 500 steps and then with conjugate gradient (CG) method, to an energy gradient of 0.005 kcal $mol^{-1} \cdot Å^{-1}$ at DFTB3/AMBER level. The resulting structure was used as initial geometry for an adiabatic mapping simulation at the same level of theory. The acylation process was described using two reaction

coordinates. The first one, Rm, described the nucleophilic attack and elimination of the leaving group, and was defined as the linear combination of the following distances: $[N_{PMA}\cdots C_{PMA}] - [C_{PMA}\cdots S_{Cys126}]$. The second coordinate, Rn, described the proton transfer from the nucleophile Cys126 to the nitrogen atom of the leaving group, and was defined as: $[S_{Cys126}\cdots H_{Cys126}] - [H_{Cys126}\cdots N_{PMA}]$. The value of Rm ranged from -1.95 to 0.60 Å, with step size of 0.15 Å, while Rn ranged from -1.60 to 2.20 Å, with step size of 0.20 Å. At each point of the grid, a DFTB3/AMBER level minimization was performed to an energy gradient of 0.005 kcal $mol^{-1}\cdot Å^{-1}$ applying a harmonic restraint of 1000 kcal $mol^{-1}\cdot Å^{-2}$ on the reference value of the reaction coordinate. From this procedure, a potential energy surface was also obtained (Figure S21). Each minimized structure obtained from the adiabatic mapping was then used as starting point for QM/MM umbrella sampling MD simulations. In this case, harmonic restraints of 100 kcal/mol were applied on the reaction coordinates. Each simulation window consisted of 45 ps of QM/MM MD equilibration, followed by 60 ps of QM/MM production. Bidimensional PMF were obtained using the WHAM 2-D approach.[19,20] The minimum free energy path on the surfaces was determined using the minimum energy pathway analysis for energy landscapes (MEPSA) software.[21]

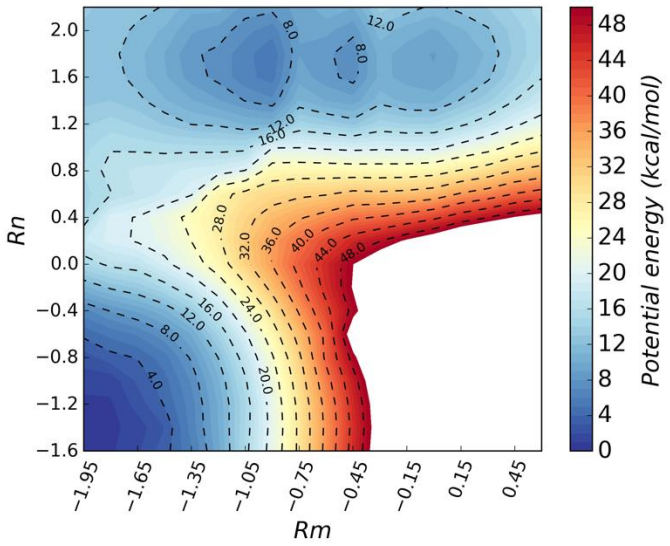


Figure S22. PES for NAAA acylation by PMA calculated at DFTB3/AMBER level. Potential energy values are given in kcal mol^{-1}, while reaction coordinates Rm and Rn are in Å.

21. PEA hydrolysis in wild-type and NAAA-overexpressing lysosomal extracts

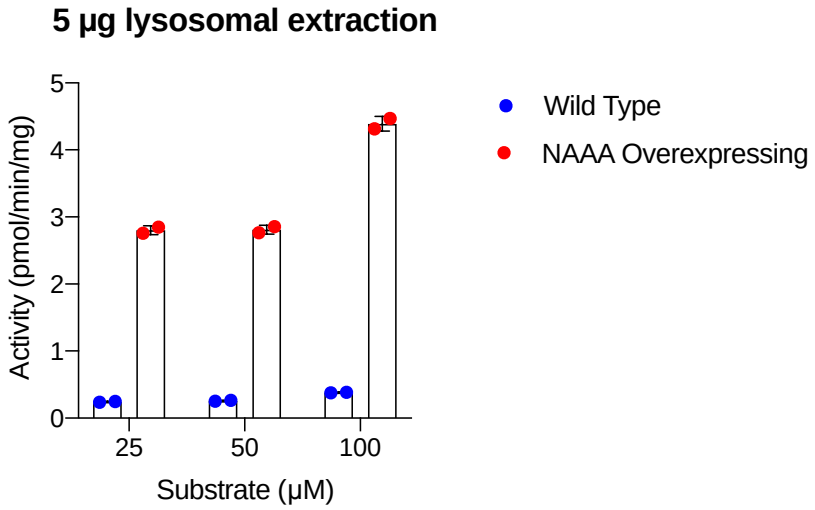


Figure S23. Hydrolysis of PEA in wild-type and NAAA-overexpressing lysosomal extracts taken from HEK-293 cells.

22. Chemical syntheses

General Informations. All chemicals were used as received unless stated otherwise. All reactions were performed under a steady overpressure of nitrogen delivered through a balloon. Anhydrous dichloromethane was directly purchased from an appropriate vendor subsequently manipulated by a syringe under a steady pressure of nitrogen. Thin-layer chromatography (TLC) analysis was conducted on HPTLC aluminum sheets (Sigma-Aldrich, silica gel 60, F254); compounds were visualized by dipping in a solution of *p*-anisaldehyde (2.5% v/v), prepared as follow: To 135 mL of absolute ethanol was added 5 mL of concentrated sulfuric acid, 1.5 mL of glacial acetic acid and 3.7 mL of *p*-anisaldehyde. The staining solution was stored in a jar covered with aluminium foil. ^{1}H NMR spectra were recorded with a Bruker DRX400 (400 MHz). Chemical shifts (δ scale) are reported in parts per million (ppm), referenced to tetramethylsilane (TMS) or the residual solvent signal. ^{1}H NMR spectra are reported in the following order: number of protons, multiplicity and approximate coupling constant (*J* value) in hertz (Hz); signals

were characterized as s (singlet), d (doublet), dd (doublet of doublets), t (triplet), dt (doublet of triplets), q (quartet), quint (quintet) m (multiplet), br (broad signal). Mass spectra were recorded on an Agilent 6410 Triple Quad LC/MS system with an ESI interface. The purity of final compounds was analyzed by analytical high-performance liquid chromatography (HPLC) employing an Agilent 6410 Triple Quad LC/MS system (Agilent Technologies., Wilmington, DE). Prior to analysis, samples were prepared in MeOH at a final concentration of 5 µg/mL. Chromatographic separation occurred on an Agilent Zorbax Eclipse XBD C18 column (2.1x50 mm, 1.8 µm particle size) by gradient elution. The flow rate was 0.4 mL/min, and the injected volume was 2.0 µL. Solvent A was water and solvent B was methanol, both additioned of 0.25% v/v acetic acid and 5mM ammonium acetate. Gradient conditions A: t(0 min): 10% A: 90% B; t(2 min): 5% A: 95% B; t(3 min): 10% A: 90% B. Purity results are presented as tR (min) and relative chemical purity (%). All tested compounds were >95% pure.

PEA. To an ice-cooled stirring solution of palmitic acid (0.186 g, 0.725 mmol, Nu-Chek Prep Inc., Elysian, MN) and triethylamine (0.25 mL, 1.813 mmol, Fisher Scientific, Waltham, MA) in Dichloromethane (Fisher Scientific, Waltham, MA), *N,N,N',N'*-Tetramethyl-*O*-(benzotriazol-1-yl)uronium tetrafluoroborate (TBTU, 0.256 g, 0.798 mmol, Sigma-Aldrich, St. Louis, MO) was added and the reaction was stirred for 60 min. Ethanolamine (0.053 g, 0.870 mmol, Sigma-Aldrich, St. Louis, MO) was added, and the mixture was stirred for 2 h. The reaction was quenched with HCl (1M), and the organic phase was washed with a saturated solution of $NaHCO_3$, then water, and brine. The organic layer was treated with Na_2SO_4, then filtered and dried under vacuum. The crude product was crystallized from heptane to give 187 mg of the desired compound. Yield = 86%. ^{1}H-NMR ($CDCl_3$ with 0.03% v/v TMS, 400 MHz) δ: 5.97 (1H, br, NH), 3.72 (2H, t, *J* = 4.8 Hz, OCH2), 3.42 (2H, q, *J* = 4.6 Hz, NCH2), 2.79 (1H, br, OH), 2.20 (2H, t, *J* = 7.4 Hz, CH2α), 1.63 (2H, quint, *J* = 7.4 Hz, CH2β), 1.40-1.15 (24H, m, (CH2)12), 0.88 (3H, t, *J* = 6.7 Hz, CH3ω); MS-ESI: calcd for C18H37NO2: 299.28, found: 300.3 [M +

H]+; tR (0.972 min). Purity 96.8%. Analytical data agree with those reported in Ottria et al, *Chem. Phys. Lipids*, **2012**, *165*, 705-711.

Palmitoylmethylamide (PMA). To an ice-cooled stirring solution of palmitic acid (0.202 g, 0.788 mmol, Nu-Chek Prep Inc., Elysian, MN) and triethylamine (0.55 mL, 3.940 mmol, Fisher Scientific, Waltham, MA) in Dichloromethane (Fisher Scientific, Waltham, MA), N,N,N',N'-Tetramethyl-O-(benzotriazol-1-yl)uronium tetrafluoroborate (TBTU 0.303 g, 0.945 mmol; Sigma-Aldrich, St. Louis, MO) was added and the reaction was stirred for 60 min. Methylamine hydrochloride (0.106 g, 1.576 mmol, Sigma-Aldrich, St. Louis, MO) was added, and the mixture was stirred for 2 h. The reaction was quenched with HCl (1M), and the organic phase was washed with a saturated solution of $NaHCO_3$, then water, and brine. The organic layer was treated with Na_2SO_4, then filtered and dried under vacuum. The crude product was crystallized from heptane to give 154 mg of the desired compound. Yield = 65%. ^{1}H-NMR ($CDCl_3$ with 0.03% v/v TMS, 400 MHz) δ: 5.39 (1H, bs, NH), 2.81 (3H, d, *J* = 4.8 Hz, NCH3), 2.16 (2H, t, *J* = 7.4 Hz, CH2α), 1.62 (2H, quint, *J* = 7.3 Hz, CH2β), 1.40-1.15 (24H, m, (CH2)12), 0.88 (3H, t, *J* = 6.7 Hz, CH3ω); MS-ESI: calcd for C17H35NO: 269.27 , found: 270.3 [M + H]+; tR (1.132 min) Purity 97.7%. Analytical data agree with those reported in Vandevoorde et al, *J.Med.Chem*, **2003**, *46*, 1440-1448.

23.References

1 Gorelik, A.; Gebai, A.; Illes, K.; Piomelli, D.; Nagar, B. Molecular mechanism of activation of the immunoregulatory amidase NAAA. *Proc. Natl. Acad. Sci. U S A* **2018**, *115*, E10032-E10040.

2 Schrödinger Release 2018-2: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2016; Impact, Schrödinger, LLC, New York, NY, 2016; Prime, Schrödinger, LLC, New York, NY, 2018.

3 Wang, J.; Zhao, L. Y.; Uyama, T.; Tsuboi, K.; Tonai, T.; Ueda N. Amino acid residues crucial in pH regulation and proteolytic activation of N-acylethanolamine-hydrolyzing acid amidase. *Biochim. Biophys. Acta* **2008**, *1781*, 710-717.

4 Olsson, M. H.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525-537

5 Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J.Y.; Wang, L.; Lupyan, D.; Dahlgren, M.K.; Knight, J.L.; Kaus, J.W.; Cerutti, D.; Krilov, G.; Jorgensen, W.L.; Abel, R.; Friesner, R.A., "OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281-296.

6 Schrödinger Release 2018-2: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2018. Maestro-Desmond interoperability tools, Schrödinger, New York, NY, 2018.

7 a) Schrödinger Release 2018-2: Glide, Schrödinger, LLC, New York, NY, 2018. b) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shaw, D. E.; Shelley, M.; Perry, J. K.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739-1749. c) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J. Med. Chem. 2004, 47, 1750-1759. d) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren,T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem*. **2006**, *49*, 6177-6196.

8 Debiec, K. T.; Cerutti, D. S.; Baker, L. R.; Gronenborn, A. M.; Case, D. A.; Chong, L. T. Further along the road less traveled: AMBER ff15ipq, an original protein force field built on a self-consistent physical model. *J. Chem. Theory Comput*. **2016**, *12*, 3926-3947.

9 Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general AMBER force field. *J. Comput. Chem*. **2004**, *25*, 1157-1174.

10 Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman P. A. AMBER 2016, University of California, San Francisco **2016**.

11 a) Schrödinger Release 2018-2: Jaguar, Schrödinger, LLC, New York, NY, 2018. b) Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. Jaguar: a high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum Chem*. **2013**, *113*, 2110-2142.

12 Marcos-Alcalde, I.; Setoain, J.; Mendieta-Moreno, J. I.; Mendieta, J.; Gómez-Puertas, P. MEPSA: minimum energy pathway analysis for energy landscapes. *Bioinformatics* **2015**, *31*, 3853-3855.

13 Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Math*. **1959**, *1*, 269-271.

14 Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in free energy space. *J. Chem. Phys*. **2007**, *126*, 054103.

15 a) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: a portable plugin for free energy calculations with molecular dynamics. *Comp. Phys. Comm*. **2009**, *180*, 1961-1972. b) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED2: New feathers for an old bird. *Comp. Phys. Comm*. **2014**, *185*, 604-613.

16 Catmull, E.; Rom, R. A class of local interpolating splines. In *Computer aided geometric design*; Barnhill, R. E., Riesenfeld, R. F. Eds.; Academic Press: New York, 1974; pp. 317–326.

17 De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of molecular dynamics and related methods in drug discovery. *J. Med. Chem*. **2016**, *59*, 4035-4061.

18 a) Schrödinger Release 2018-2: Glide, Schrödinger, LLC, New York, NY, 2018. b) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shaw, D. E.; Shelley, M.; Perry, J. K.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem*. **2004**, *47*, 1739-1749. c) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem*. **2004**, *47*, 1750-1759. d) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren,T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177-6196.

19 a) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem*. **1992**, *13*, 1011-1021. b) Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun*. **1995**, *91*, 275-282. c) Kumar, S.; Rosenberg, J. M.; Bouzida, D.;

Swendsen, R. H.; Kollman, P. A. Multidimensional free-energy calculations using the weighted histogram analysis method. *J. Comput. Chem*. **1995**, *16*, 1339-1350.

20 Grossfield, A. "WHAM: the weighted histogram analysis method", version 2.0.9, http://membrane.urmc.rochester.edu/wordpress/?page_id=126

21 Marcos-Alcalde, I.; Setoain, J.; Mendieta-Moreno, J. I.; Mendieta, J.; Gómez-Puertas, P. MEPSA: minimum energy pathway analysis for energy landscapes. *Bioinformatics* **2015**, *31*, 3853-3855.